ADDUCTOR SPASMODIC DYSPHONIA VERSUS MUSCLE TENSION DYSPHONIA: EXPLORING THE PRECISION OF PHONATORY BREAK ANALYSIS AS A DIAGNOSTIC TEST

by

Melissa Whitchurch

A thesis submitted to the faculty of
The University of Utah
in partial fulfillment of the requirements for the degree of

Master of Science

in

Speech-Language Pathology

Department of Communication Sciences and Disorders

The University of Utah

August 2008

THE UNIVERSITY OF UTAH GRADUATE SCHOOL

SUPERVISORY COMMITTEE APPROVAL

of a thesis submitted by

Melissa Whitchurch

This thesis has been read by each member of the following supervisory committee and by majority vote has been found to be satisfactory.

Chair: Nelson Roy

Sean Redmond

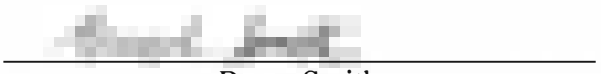
THE UNIVERSITY OF UTAH GRADUATE SCHOOL

FINAL READING APPROVAL

To the Graduate Council of the University of Utah:

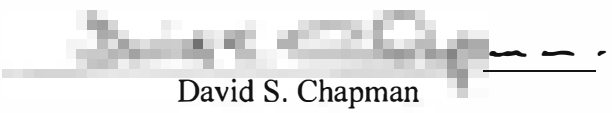
I have read the thesis of Melissa Whitchurch in its final form and have found that (1) its format, citations, and bibliographic style are consistent and acceptable; (2) its illustrative materials including figures, tables, and charts are in place; and (3) the final manuscript is satisfactory to the supervisory committee and is ready for submission to The Graduate School.

Approved for the Major Department



Bruce Smith
Chair/Dean

Approved for the Graduate Council



David S. Chapman
Dean of The Graduate School

ABSTRACT

Adductor spasmodic dysphonia (ADSD) and muscle tension dysphonia (MTD) have voice characteristics that can mimic each other, thus leading to diagnostic confusion. Given the perceptual similarity, specific distinguishing features are needed to objectively compare characteristics of ADSD and MTD. Intraword phonatory breaks, comprising of a complete absence of phonation, are one such acoustic feature that needs further research. This investigation assessed the diagnostic worth of acoustic analysis of phonatory breaks as a possible objective test to distinguish ADSD from MTD.

Fifty-nine subjects with MTD and 41 subjects with ADSD were recorded reading an all-voiced consonant sentence: *Early one morning a man and a woman were ambling along a one-mile lane running near Rainy Island Avenue.* The presence and duration in milliseconds of any within-word phonatory breaks were measured. Estimates of sensitivity, specificity, positive and negative predictive value, and likelihood ratios were calculated to determine the precision and worth of phonatory break analysis as a clinical diagnostic test.

Results revealed that (a) individuals with ADSD showed a higher mean number of phonatory breaks as compared to individuals with MTD; (b) phonatory breaks occurred in participants with ADSD, particularly at durations less than 60 ms as compared to the MTD group. Also, the ideal duration of phonatory break measurement was 40 ms; (c) all measures of diagnostic precision were markedly better in males. Males with MTD rarely

evidenced phonatory breaks, and no male subject with MTD had a break greater than 70 ms; (d) as the number of phonatory breaks increased, diagnostic precision also increased. For example, when a patient has more than four phonatory breaks, it can be quite confidently concluded that the patient has ADSD. It can be concluded that combining information regarding duration and frequency, along with knowledge of gender, improved diagnostic test performance. Automation of the acoustic analysis procedure should be explored.

TABLE OF CONTENTS

LIST OF FIGURES

Figure Page

LIST OF TABLES

Table Page

ACKNOWLEDGMENTS

One of the benefits of attending the University of Utah Speech-Language Pathology program was its "thesis-optional" requirement. From the beginning I fully intended to opt out of the thesis alternative. However, there was something about the way Dr. Roy presented the option of doing a thesis that changed my mind. It is a decision I will never regret. I could not have asked for a better supervisor and tutor throughout this process. Dr. Roy was patient in guiding my research and writing, teaching me lessons I could not have learned otherwise. I also appreciate the time and energy my committee members, Dr. Blomgren and Dr. Redmond, offered. Thanks also goes to Dr. Merrill, who completed numerous statistical analysis. Most of all, I could not have made it through my graduate experience without my trusted study partner, Lindsay Downs. She understood better than anyone the doubts and concerns that arose over the past 2 years. She stood beside me as a dedicated friend and cheerleader to the very last edit, print, and defense. Lindsay not only completed the reliability for my thesis willingly and without complaint but listened as I shared frustrations and successes along the way. Without doubt, I owe the completion of my graduate degree to her never-ending support. Additionally, I must recognize the support and encouragement of my parents. They provided consistent faith in my efforts and capability to complete this task. They never wavered in believing my ability would be sufficient and showed glowing pride at my graduation. Mom and Dad have never put limits on the things I can accomplish and continue to show confidence in my life decisions. Thank you to the rest of my family and friends who also showed positive support in my educational endeavors.

INTRODUCTION

A near consensus exists that spasmodic dysphonia is a neurogenic, action-induced, focal laryngeal dystonia with several subtypes. Of these subtypes, adductor spasmodic dysphonia (ADSD) is the most common form wherein the speaker experiences intermittent or continuous hyperadduction (overclosure) of the true vocal folds during speech (Cannito & Woodson, 2000). This involuntary hyperadduction has been associated with spasmodic muscle bursts in the laryngeal adductor muscles, contributing to debilitating voice breaks and an effortful, strained-strangled voice quality (Ludlow, 1995). Despite recent advances, the precise etiology of this enigmatic disorder remains unknown and no definitive lesion loci or specific pathological processes can be identified in the majority of cases (Cimino-Knight & Sapienza, 2001; Roy, Mauszycki, Gouse, & Smith, 2007). Currently, the diagnosis of ADSD is based almost exclusively on auditory-perceptual features (Langeveld, Drost, Frijns, Zwinderman, & Baatenburg de Jong, 2000). However another enigmatic voice disorder—muscle tension dysphonia (MTD)—can mimic the voice characteristics of ADSD, thus leading to diagnostic confusion and possibly inappropriate management (Roy, Gouse, Mauszycki, Merrill, & Marshall, 2005; Roy, Smith, Allen, & Merrill, 2007). Unlike ADSD, MTD is considered a "functional" voice disorder wherein excess or dysregulated laryngeal and extralaryngeal muscle activity can create a strained, effortful voice quality similar to ADSD. Poorly regulated

activity of the intrinsic and extrinsic laryngeal muscles is cited as the proximal cause of MTD, but the origin of this muscle activity is often unclear and has been attributed to multiple sources including psychological and personality factors (Roy, Mauszycki et al., 2007).

Given the perceptual similarity of both MTD and ADSD and the potential for misclassification even among experienced clinicians, there is a need to objectively compare characteristics of ADSD and MTD to determine if there are features that reliably distinguish the two disorders (Roy, Mauszycki et al., 2007). In the following section, the extant literature comparing ADSD and MTD is reviewed.

ADSD Versus MTD

Only a handful of studies have directly compared ADSD with MTD to identify markers that could potentially distinguish the two disorders. A few differences have been observed during fiberoptic laryngoscopy and phonatory airflow measurement (Higgins, Chait, & Schulte, 1999; Roy, Mauszycki et al., 2007). Leonard and Kendall (1999) concluded that transnasal fiberoptic endoscopy alone *does not* offer a reliable means of differential diagnosis. Higgins and coauthors also reported differences in phonatory airflow characteristics between groups of women with ADSD and MTD. Although these investigations identified some differences between the groups, the authors did not provide any information in the form of sensitivity or specificity estimates regarding the precision of these measures as potential diagnostic markers of ADSD or MTD.

In contrast, Roy and colleagues (2005) recently assessed the task-specificity hypothesis in ADSD, and whether task-specificity might serve to distinguish ADSD from

MTD. The task-specificity hypothesis states that sign expression in ADSD is task dependent, with certain vocal tasks provoking more frequent and severe phonatory signs than other tasks. The authors employed listener ratings to confirm that the dysphonia associated with sustained vowels was perceived to be less severe than during connected speech in ADSD whereas patients with MTD did not show variable performance based on voice context. In light of respectable sensitivity and specificity estimates, the authors concluded that task-specificity, as evidenced by differential performance during sustained vowels versus connected speech, was a useful diagnostic marker for ADSD and advised that auditory-perceptual comparison of these two voice contexts during routine clinical assessment should improve the likelihood of correct differential diagnosis.

More recently, Roy, Mauszycki et al. (2007) also examined the value of task specificity in distinguishing ADSD (n=29) from MTD (n=33). However, in this investigation the researchers compared dysphonia severity ratings for a sentence loaded with voiced consonants versus a sentence loaded with voiceless consonants. The all-voiced sentence was *Early one morning a man and a woman were ambling along a one-mile lane running near Rainy Island Avenue.* In contrast, the sentence laden with voiceless consonants was *He saw half a shape mystically cross fifty or sixty steps in front of his sister Kathy's house.* Five listeners rated dysphonia severity for both contexts on a 10-cm visual analog scale ranging from normal to profoundly abnormal. Results for the ADSD patients showed a significant difference based on phonetic context, with the all-voiced sentence rated as significantly more severe as compared to the voiceless-laden sentence. For patients with MTD, no difference was observed between the two sentence types. Like the previous Roy et al. (2005) study, this study provided sensitivity and

specificity measures with a receiver operating characteristic (ROC) curve to help interpret the results. The results confirmed that differences in sign expression based on phonetic (i.e., voicing) context provided excellent specificity which ranged from 90-100% depending upon the cutoff value used, but only 48% sensitivity. The authors concluded that further research to identify additional diagnostic markers and tests is needed to improve diagnostic precision.

Roy, Smith, et al. (2007) also examined the value of lidocaine block of the recurrent laryngeal nerve (RLN) as a diagnostic test to distinguish ADSD and MTD. Dedo and Shipp (1980) previously reported using RLN block in patients with ADSD to determine their candidacy for a surgical treatment known as RLN sectioning. Patients with a positive response, that is, reduced dysphonia severity during the block condition, were judged by Dedo and Shipp to be good candidates for surgical sectioning of the RLN as a treatment for their ADSD symptoms. Therefore, Roy and colleagues (2007) reasoned that positive response to RLN lidocaine block might also be used to potentially distinguish ADSD from MTD. To test this hypothesis, 23 patients with ADSD and 20 patients with MTD underwent unilateral lidocaine block. Each patient read an all-voiced sentence before the nerve block and 10 min after complete paralysis of the right vocal fold due to injection of lidocaine into the vicinity of the right RLN. Each participant's voice was self-rated based on severity, vocal effort, and laryngeal tightness before and after the block. In addition, six listeners made judgments regarding severity, breathiness, and strain using a 100 mm visual analog scale. Interestingly, both groups responded favorably to the RLN block procedure, and no significant differences between the groups were found on any of the patient- or listener-based ratings. Based upon extremely poor

sensitivity and specificity estimates as well as disappointing ROC curves, the authors concluded that patient response to the RLN block is a worthless diagnostic test and should not be used to differentiate MTD and ADSD.

In addition to the research of Roy and colleagues—based principally on auditory-perceptual judgments—Sapienza, Walton, and Murry (2000) used acoustic analysis techniques to evaluate the acoustic characteristics of ADSD and to determine whether differences exist between ADSD and MTD. Three particular acoustic measures have received attention and have been repeatedly employed in a number of their studies: phonatory breaks, aperiodicity, and frequency shifts. Phonatory breaks were defined as an absence of voicing of at least 50-ms within words. A frequency shift was defined as a 50 Hz change in fundamental frequency over a 50-ms period. Aperiodicity was defined as cycle-to-cycle variability of the fundamental period as evidenced by irregular period duration or non-repetitive cycles (Sapienza, Murray, & Brown, 1998).

Sapienza and colleagues (1998) used these acoustic measures to compare individuals with normal voices (n=14) and individuals diagnosed with ADSD (n=14). Although recordings of the sustained vowel /a/, repeated /pa/, paragraph recitation, and picture description were recorded, only the results of the sustained vowel production were reported. Vowel duration varied according to the ability of the subject and ranged from 2.8 to 21.7 s. Within the sustained vowel context, the results indicated that individuals with normal voices had no phonatory breaks, frequency shifts, or aperiodic segments that met stated criteria. Out of the 14 individuals with ADSD, only 9 produced phonatory breaks (i.e., 64.3%). The mean number of phonatory breaks for all individuals with ADSD was 10.1. It should be noted that the percentage of phonatory breaks

constituting the sustained vowel was less than 1% for half the subjects, although the percentage of phonatory breaks in the sustained vowel context ranged from 0% to 28%. The majority (81%) of the phonatory breaks occurred during the midpoint of the vowel. The number, duration, and percentages of frequency shifts and aperiodic segments will not be reviewed in detail because as the results are only peripherally germane to the purpose of the current study. Sapienza concluded that frequency shifts, aperiodicity, phonatory breaks, and normal phonation are all acoustic characteristics of ADSD in sustained vowel production. These investigators asserted that these parameters can be identified manually with relative ease, help to characterize ADSD, and distinguish ADSD from normal voices.

In a follow-up investigation, Sapienza, Walton, and Murry (1999) sought to assess the frequency and duration of such acoustic events, comparing subjects with ADSD to normal subjects across two tasks: sustained vowel production and reading. The experimenters also examined whether patients with ADSD showed similar acoustic features in both sustained vowel and reading tasks. In the study, 14 women with ADSD and 14 age-matched vocally normal women provided the voice samples. It was reported that the women with ADSD displayed a higher frequency of aberrant acoustic phenomena during both vocal tasks as compared to the control group. When comparing vowel production with reading tasks, however, ADSD patients had more phonatory breaks, frequency shifts, and aperiodic segments during vowel production. Aperiodic segments were the most prominent feature. This result is at odds with Roy and colleagues' later finding that in ADSD sustained vowels were perceptually less symptomatic when compared to connected speech. This difference in results may be related to differences in

analysis methods (i.e., auditory-perceptual vs. acoustic). The authors (Sapienza et al.) concluded that in ADSD there was more acoustic variability during sustained vowels as compared to reading, although both tasks contain valuable information and should be used in assessment.

Finally, Sapienza and colleagues (2000) attempted to identify differences in acoustic measures in the voices of patients with ADSD versus MTD. The investigators acoustically analyzed speech samples of sustained vowel production and the Rainbow Passage. Participants consisted of 10 individuals with ADSD and 10 individuals with MTD. The purpose of this study was to help differentiate ADSD from MTD using specific noninvasive acoustic criteria. Consistent with the previous Sapienza et al. (1998) study, the investigators looked at three acoustic phenomena: phonatory breaks, aperiodicity, and frequency shifts during either sustained vowel production or reading. The same definitions for phonatory breaks, aperiodicity, and frequency shifts were employed as in the previously reported Sapienza et al. (1998, 1999) studies.

Each voiced segment only received one classification. All individuals with ADSD produced at least one phonatory break (ranging from 1 to 50 breaks) during sustained vowel production. Patients with MTD produced no phonatory breaks. During reading, only four breaks were identified for all 10 patients with ADSD, and again patients with MTD evidenced no phonatory breaks. It should be noted, however, that the absence of phonatory breaks in the MTD group likely reflects inclusion/exclusion criteria employed by the researchers. That is, the researchers excluded any individuals with MTD if they had perceptual evidence of voice breaks during sustained vowel productions; thus, it is not surprising that there was no acoustic evidence of phonatory breaks in the MTD group.

Thus, from the experimenters' exclusion criteria it is difficult to determine whether phonatory breaks are in fact pathognomonic of ADSD and serve to reliably distinguish ADSD from MTD. Furthermore, Sapienza and colleagues' studies provided no estimates of sensitivity or specificity or any other measures of diagnostic precision. Therefore, the diagnostic value of measures such as the presence, frequency and duration of phonatory breaks to distinguish ADSD and MTD is unknown.

Recently, Rees, Blalock, Kemp, Halum, and Koufman (2007) compared ADSD and MTD using spectral analysis. The authors asserted that MTD is a voice disorder secondary to occult glottal incompetence and that spectral analysis can help to differentiate between ADSD and MTD. Spectography displays a graphic image of frequency and amplitude as a function of time. Thus, spectography relies on subjective visual interpretation of frequency domain acoustic data. In the study, spectrograms were created from 2-s speech samples taken from the Rainbow Passage (Fairbanks, 1960). Two raters experienced in spectrographic interpretation assigned a diagnosis of ADSD or MTD based on presence or absence of certain spectrographic features. Rees et al. concluded that the following spectral features distinguish ADSD from MTD: (a) abrupt voice breaks, (b) irregular wide-spaced vertical striations, and (c) formant definition preservation, and provided near perfect measures of sensitivity and specificity.

The authors concluded that spectrographic analysis of voice samples is an effective means of differential diagnosis. However, serious methodological problems exist with this study, which ultimately discredit the results and render the conclusions invalid. The number of problems are too numerous to review in detail, but the following is a brief synopsis. First, no evidence of baseline equivalence in dysphonia severity

between the groups was provided. Thus it is not known whether the spectrographic features which distinguished the two groups merely reflect differences in severity between the groups at baseline, rather than true phenomenological differences between the groups independent of severity. Second, the original two-second spectrogram from the patient's chart, which was used to render the diagnosis, was retrieved from the original files, and the same speech-language pathologist (SLP) who acquired and interpreted the original spectrogram—and ultimately rendered the diagnosis—was then asked to reinterpret the same spectrograms and provide a differential diagnosis. This is merely a measure of inter-judge reliability, not an evaluation of the precision of a diagnostic test. The fact that the SLP rendered the same diagnosis says nothing about the validity of the diagnostic test, and the impressive sensitivity and specificity measures reported merely reflect internal consistency, rather than an adequate assessment of the diagnostic precision of the test (i.e., as compared to some independent standard). In this regard, the authors provided no gold standard diagnostic criteria for MTD or ADSD to verify that the correct diagnosis had ever been made. Aside from the spectrographic data apparently used in the original exam, there was no evidence offered that authors actually made the correct diagnosis. That is, the same test used to make the original diagnosis was merely repeated. Third, the 2-s spectrographic sample did not contain identical content across subjects and was presumably included originally in the chart, because the spectrographic sample best exemplified the spectral features of interest. Taking the ideal 2-s segment from a 35-s passage hardly represents a fair assessment of a test's performance. Fourth, the majority of MTD patients included in the study also had underlying glottic insufficiency, including vocal fold paresis. Thus these are patients

with secondary MTD, not primary MTD. This unnecessarily complicates interpretation of the results and limits the generalizability of the tests worth to primary MTD. To complicate matters, prior to acquisition of the spectrographic data, the authors administered voice therapy to the MTD patients, apparently to remove compensatory laryngeal behaviors. This process included increasing breath support, softening glottal attacks, and reducing laryngeal and neck muscle tension. No mention was made as to whether any of ADSD patients underwent such therapy. The logic and necessity of applying voice therapy prior to administering the diagnostic test (i.e., generating the spectrogram) make little sense and again do not permit an accurate assessment of the test's clinical value without administering voice therapy. Finally, using spectography in this manner merely represents a subjective visual-perceptual judgment, based upon training and experience of the interpreter. A second rater trained by the first also rated the samples; however no estimates of inter-rater reliability were provided. The two raters who participated had 9 and 20 years experience using spectral analysis. It is difficult to ascertain whether any SLP could be taught to use this analysis approach for the purpose of differential diagnosis of ADSD or MTD. Unless reliability within and across judges for this visual-perceptual rating task can be established, it likely remains a clinically impractical test.

From the previous literature review it is evident that there remains a critical need to identify distinguishing characteristics and improve differential diagnosis for ADSD and MTD.

Purpose of the Study

The results of the literature review suggest that further study is needed to identify diagnostic features or tests that will reliably distinguish ADSD and MTD. In this regard, phonatory breaks are often considered pathognomonic of ADSD, but little is known regarding the actual frequency of occurrence of this feature in ADSD and, more importantly, whether patients with MTD also share this acoustic/perceptual feature. To determine the diagnostic value of phonatory breaks this study addressed the following questions: Do subjects with ADSD and MTD demonstrate acoustic evidence of phonatory breaks? Are phonatory breaks specific to ADSD? As indexed by estimates of sensitivity, specificity, positive predictive values, negative predictive values, and likelihood ratios, what is the diagnostic precision of phonatory breaks as a marker of ADSD that distinguishes it from MTD?

METHODS

Participants

An existing database of archived recordings was used for this research with University of Utah IRB approval. Pretreatment voice samples from 59 subjects with MTD (10 males, mean age, m = 49.20 years, standard deviation, SD = 15.94 years; 49 females, m = 48.47, SD = 15.76) and 41 subjects with ADSD (19 males, m = 47.42 years, SD = 11.96 years; 22 females, m = 50.00, SD = 13.99) were used for analysis. The female-to-male ratio among participants within each voice disorder was consistent with proportions described elsewhere in the literature (Ludlow & Mann, 2003; Roy, Bless, Heisey, & Ford, 1997). All patients in the database were diagnosed with either MTD or ADSD after a comprehensive speech and voice evaluation, including videolaryngoscopy by both an otolargyngologist and a speech-language pathologist who specialized in voice disorders and had extensive experience in voice disorder assessment and management. In both ADSD and MTD, videolaryngostroboscopy confirmed a structurally normal larynx, free of mucosal disease and/or frank vocal fold paresis or paralysis.

A provisional diagnosis of ADSD was offered following the guidelines described by Cannito and Kondraske (1990). In brief, patients were considered as presenting with ADSD if they demonstrated the following: (a) absence of perceptual symptoms of the classical dysarthrias, (b) auditory-perceptual characteristics consistent with the disorder (continuous or intermittent strained-strangled voice quality and no obvious tremor during

phonation), (c) occasional moments of normal sounding voice, (d) improved voice for nonspeech vocalizations, and (e) improved voice quality for phonation at high pitch. Finally, the provisional diagnosis of ADSD was later corroborated by an unsuccessful trial of voice therapy conducted by an experienced voice clinician. The participants with ADSD had not previously received intrafold botulinum toxin injection(s) at the time of audio recording. However, all participants with ADSD later responded favorably to periodic botulinum toxin injections into the laryngeal adductor muscles (based upon patient self-assessment, SLP, and physician report). Admittedly, in the absence of a gold standard for diagnosis, the foregoing inclusion and exclusion criteria improve the confidence of the ADSD diagnosis but do not completely exclude the possibility of mixed SD (i.e., both ADSD and another form of spasmodic dysphonia known as abductor spasmodic dysphonia, which is characterized by intermittent breathiness and spasmodic abduction of the vocal folds) or SD plus MTD.

In contrast, participants with MTD demonstrated dysphonia (not aphonia) that was characterized by myriad voice qualities. In addition, they also reported additional features considered stereotypic of excess perilaryngeal tension, such as significant discomfort during focal palpation of the hyolaryngeal sling, and a reduced thyrohyoid space suggesting laryngeal elevation (Aronson, 1990; Higgins, et al., 1999; Roy, Bless, Heisey, & Ford, 1997; Roy & Leeper, 1993). The sustained improvement with short-term behavioral intervention served to substantiate the diagnosis of MTD and excluded ADSD as a possible diagnostic alternative.

Procedures

Voice Stimuli

Each participant was recorded reading an all-voiced consonant sentence: *Early one morning a man and a woman were ambling along a one-mile lane running near Rainy Island Avenue* (Dedo & Shipp, 1980). This all-voiced context has been shown to maximally provoke dysphonic symptoms in ADSD (Roy, Mauszycki, et al., 2007). Recordings were made with research-quality audio recording equipment in a quiet room. The sentences were originally digitized using Kay Elemetrics Multi-Speech, Computerized Speech Laboratory, CSL, Model 3700. Following digitization the recordings were viewed as an acoustic waveform using Multi-Speech, version 3.1.7, (KAYPentax Corp., Lincoln Park, NJ).

Baseline Equivalence Testing

To permit a valid test of phonatory breaks as a potential diagnostic marker, it was necessary to first establish that the two disorder groups were equivalent on overall dysphonia severity. If the two groups shared similar levels of dysphonia severity, valid comparisons (i.e., statistical tests) of between-group differences on number and duration of phonatory breaks would be permitted, and would not be confounded by unequal levels of dysphonia severity.

Auditory-perceptual data derived from a previously related investigation were employed to establish equivalence in dysphonia severity (Houtz, 2006). In a study by Houtz, an auditory-perceptual rating task was conducted using the identical subject pool and an all-voiced sentence. Houtz conducted the auditory-perceptual rating task as

follows: voice stimuli (normalized for amplitude) were transferred from *Multi-Speech*, version 2.7 to *Alvin*, version 1.01, public domain software designed for various listening experiments (simple labeling, judgments of voice quality, discrimination, sentence intelligibility, etc. [Hillenbrand & Gayvert, 2005]). Levels were adjusted using the normalized audio files so that all samples would be presented at comparable loudness levels for the listening task. Five SLP graduate students who had completed graduate coursework in the assessment and management of voice disorders were selected to be raters. Each listener rated randomly presented voice samples by "clicking" a mouse connected to a cursor on a visual analog scale (VAS), ranging from "normal voice" on the extreme left of the scale to "profoundly abnormal voice" on the extreme right. The software program then translated each analog scale rating to a number ranging from 0 to 1000, with larger numbers corresponding to increasing dysphonia severity. The listeners were also blinded to the specific diagnostic categories and to the purpose of the study.

Inter- and intra-rater reliability estimates of perceptual judgments of severity were calculated and confirmed acceptable listener reliability (see Houtz, 2006, for details) and permitted the analysis of baseline equivalence which follows. An independent samples *t* test was used to compare mean ratings of the all-voiced sentence from the listener ratings for the ADSD and MTD groups. Among males, the mean severity rating at baseline was 470.4 (SD = 250.1) for the ADSD group (n = 19) and did not differ significantly from the mean of 283.5 (SD = 298.2) for MTD group (n = 10) (t [27] =1.79, p = .0845). Among females, the mean severity rating at baseline was 493.6 (SD = 310) for ADSD (n = 22) and 522.5 (SD = 294.9) for MTD (n = 49). These means were also not significantly different (t [69] = -0.38, p = .7081). Thus, the results from Houtz confirmed

that each group was equivalent on dysphonia severity at baseline, which would permit a valid test of the worth of phonatory break analysis as a diagnostic test.

Number and Duration of Phonatory Breaks

With the all-voiced sentence displayed as an acoustic waveform in Multi-Speech, version 3.1.7, (KAYPentax Corp., Lincoln Park, NJ), the presence and duration of any within-word phonatory breaks were measured. A phonatory break consisted of a complete absence of phonation. The absolute duration of the phonatory break was measured in milliseconds by marking the initiation of the break with the last clear positive glottal pulse and termination of the phonatory break with the first clear positive glottal pulse at the return of phonation. More than one phonatory break was possible within each word. If the participant began the word, experienced a phonatory break, and reinitiated the word, the phonatory break was not counted. When ambiguity arose regarding the presence or absence of a voice break, the presence of a phonatory break was confirmed using auditory-perceptual evaluation combined with a wide-band spectrogram to determine the presence/absence of a voice bar. Occasionally it was perceptually apparent that a phonatory break occurred at the initiation of a word. However, this phonatory break was not counted unless there was acoustic evidence of prior initiation of voicing followed by a break in phonation. The total number of breaks for each participant was manually counted by the experimenter; the location (i.e., the individual word) where phonatory breaks occurred was also recorded, as well as the duration of each break. It should be noted that Sapienza and colleagues arbitrarily defined a phonatory break as any intraword interruption in phonation greater than 50-ms. However, the experimenter was interested in exploring whether other break durations

might be more informative of differences between the two disorder groups. Thus, all intra-word phonatory breaks greater than 20-ms were measured (see Figure 1 for an example of a phonatory break from an individual ADSD).

Reliability

Fifteen percent of the voice samples of ADSD and MTD were randomly selected for remeasurement to determine both intrajudge and interjudge reliability. Each sample was re-measured for number and duration of phonatory breaks by the original experimenter and another trained graduate student. Mean number and duration of phonatory breaks for ADSD and MTD were calculated individually and within groups as a whole. To determine inter- and intra-judge reliability for the measurement of number and duration of phonatory breaks Pearson product moment correlation coefficients (r) were computed. Pearson correlation coefficients of $r = 0.84$ ($p = 0.01$) and $r = 0.98$ ($p = 0.01$) for number of phonatory breaks and duration of breaks respectively, indicated acceptable interjudge reliability. For correlation of intrajudge reliability, $r = 0.98$ ($p = 0.01$) for the number

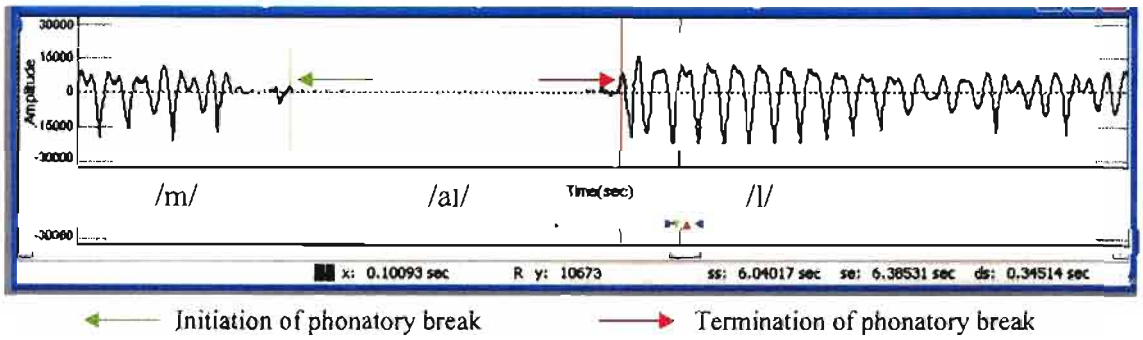


Figure 1. Phonatory break from ADSD patient on the word *mile* measuring 100.93-ms

of phonatory breaks and $r = 0.99$ ($p = 0.01$) for duration of phonatory breaks. These correlation coefficients also confirm acceptable levels intrajudge reliability.

Description of Statistical Analyses

A series of statistical analyses were undertaken to determine (a) whether differences existed between the groups (ADSD vs. MTD) on the presence, number, and duration of phonatory breaks; (b) whether differences in phonatory breaks existed between males and females in each group; and (c) the value of phonatory break analysis as a potential tool in differential diagnosis. In this regard, several indices of diagnostic test performance were calculated. For example, sensitivity and specificity estimates were calculated as indices of the diagnostic precision of phonatory breaks. Sensitivity is the proportion of subjects with the disease (i.e., cases) who have a positive test, whereas the specificity is the proportion of subjects without the disease (i.e., noncases) who have a negative test. In tests that yield continuous data like those produced in this study, several values of sensitivity and specificity are possible, depending on the cutoff point chosen to define a positive test. This trade-off between sensitivity and specificity can also be displayed graphically using a ROC curve. To generate a ROC curve, the investigator selects several cutoff points and determines the sensitivity and specificity at each point. Sensitivity (or the true positive rate) is plotted on the Y-axis as a function of 1-specificity (the false positive rate) on the X-axis. An optimal diagnostic test is one that reaches the upper left corner of the graph. A worthless test follows the diagonal from the lower left to the upper right corners, suggesting that at any cutoff the true-positive rate is the same as the false-positive rate. Estimates of sensitivity, specificity, and the ROC curve provide

valuable information regarding the strength of the presence of phonatory breaks as diagnostic marker/test.

Four additional indices of diagnostic precision were also calculated: positive predictive value (PV+), negative predictive value (PV-), positive likelihood ratio (LR+), and negative likelihood ratio (LR-). Figure 2 has a summary chart of statistical analysis and formulas. PV+ and PV- were used to determine diagnostic precision rates of those with a positive or negative test (i.e., those with and without phonatory breaks). PV+ and PV- adjust for differential base rates of the disorder states in the sample (i.e., unequal subject numbers in the two groups). PV+ reflects the proportion of cases among individuals with positive results who will have the disorder of interest, whereas PV- reflects the proportion of individuals who have a negative test and do not have the disorder of interest. The PV+ of a test can be determined by calculating the percentage of true positives/(true positives + false positives). The PV- can be determined by calculating the percentage of true negatives/(true negatives + false negatives). Because they are proportions, predictive values range from 0 to 1. Within the context of ADSD and MTD,

		Reference Standard	
		"Have ADSD"	**"Have MTD"**
Test Outcome	**Positive Test** (i.e., Evidence of phonatory breaks)	**A** *True positives*	**B** *False Positives*
	Negative Test (i.e., No evidence of phonatory breaks)	**C** *False negatives*	**D** *True negatives*

Figure 2. Graphic representation of statistical analysis.

Sensitivity = A / (A + C)
Specificity = D / (B + D)
PV+ = A / (A + B)
PV– = D / (C + D)
LR+ = sensitivity / (1 – specificity)
LR– = (1 – sensitivity) / specificity

if a test yields a high PV+, this indicates a high probability that the person who has a positive test (i.e., shows acoustic evidence of phonatory breaks) will be classified as having ADSD. In contrast, a high PV- indicates a high probability that the person who has a negative test (i.e., shows no acoustic evidence of phonatory breaks) will be classified as not having ADSD, but alternatively will have MTD. These tests further define diagnostic precision of phonatory breaks in distinguishing ADSD from MTD.

Likelihood ratios also provide additional information about the value of a diagnostic test and help diminish problems with sensitivity, specificity, PV+, and PV- and are not dependent upon base rate differences among the disorders in the sample. Thus, although the results of sensitivity, specificity, PV+, and PV- often lead to similar conclusions, likelihood ratios allow more specific interpretation. The likelihood ratio incorporates both the sensitivity and specificity of the test and provides a direct estimate of how much a test result will change the odds of having a disease. The likelihood ratio for a positive result (LR+) yields information regarding how much the odds of the disease increase when the test is positive. Specifically, LR+ is calculated by determining the ratio of true positive cases (sensitivity) to false positive cases (1-specificity) and gives information regarding the likelihood that an individual has a particular disorder. When LR+ yields a number greater than 10, the value of the diagnostic test is "high." If the LR+ yields a value of 3, there is a "moderate" likelihood that the test suggests the person has the disorder, but is not conclusive and therefore should be interpreted with caution. If the test yields a LR+ of 1, the diagnostic test does not help to diagnose a specific disorder. LR- produces an estimate that helps determine whether an individual does not have a particular disorder when the diagnostic test does not identify them as such. LR-

gives information regarding how much the odds of the disease decrease when a test is negative. It is calculated by determining the ratio of false negative cases (1-sensitivity) to true negative cases (specificity). Therefore, when a LR- yields number of less than 0.10, one can be confident that the diagnostic test helps to rule out the presence of a disorder. If LR- yields a score of less than 0.30, there is a "moderate" likelihood that the negative test score helps rule out the disorder but it should be interpreted with caution due to variability. If the LR- yields a score of 1, the diagnostic test does not help to eliminate a specific disorder. Confidence intervals (CI) are paired with both LR+ and LR- to give a range of values surrounding the observed value (Dollaghan, 2007).

RESULTS

The ADSD and MTD groups contained unequal numbers of males and females; thus a chi-square analysis was computed to compare the proportions of males to females in each group and revealed that differences between the number of men and women in the two groups were significant. That is, participants in the ADSD group had a statistically significant lower percentage of women as compared to the MTD group 54% vs. 83%; X^2 1 = 10.15, p = <0.001. This finding is consistent with gender differences reported elsewhere in the literature regarding the proportions of males and females in each disorder group. However, given the unequal proportions of males to females in the two groups, all later statistical analyses are reported separately for males and females, or where appropriate, results were adjusted for gender as a possible covariate. In cases where gender was not a significant variable, and/or for completeness, results are also reported for the combined males and females for each group.

Group Differences

To assess whether significant differences existed between the groups for number and duration of phonatory breaks, a series of independent samples *t* tests were conducted to compare the mean number of phonatory breaks between ADSD and MTD using 10-s intervals ranging from 20 to 100-ms. Significant differences emerged between groups at 20, 30, 40, 50, and 60-ms (see Table 1). As the duration of the phonatory breaks increased beyond 70-ms, no significant between-group differences were observed.

Interaction terms involving gender and ADSD versus MTD status were not statistically significant in any of the models considered in Table 1. That is, the higher mean number of breaks among ADSD compared with MTD patients did not significantly differ between males and females.

Tables 2 and 3 compare the proportions of participants with ADSD and MTD according to gender and duration of phonatory breaks. Chi-square comparisons were computed, and associated *p* values are reported. Inspection of Tables 2 and 3 reveals that as compared to females with MTD, males with MTD have relatively fewer phonatory breaks at all durations, with no evidence of breaks occurring of >70-ms. Differences between the percentages of male participants with ADSD versus MTD were significant at 40-ms and 70-ms at the 0.05 level (Table 2), whereas proportions of female participants with MTD and ADSD did not differ significantly at any duration (Table 3).

Accuracy of Phonatory Break Analysis as a Diagnostic Test

Sensitivity, specificity, PV+, PV-, LR+, and LR- were calculated to assess the diagnostic precision of phonatory breaks for correct identification of ADSD (i.e., sensitivity) versus correct identification of MTD (i.e., specificity). Table 4 displays the diagnostic performance data using duration of a phonatory break as the primary cutoff criterion for both males and females. The two most notable durations of phonatory breaks were found at 40-ms and 70-ms. Using the 40-ms cutoff produced a sensitivity value of 66%, (i.e., 66% of subjects with ADSD were correctly classified as having ADSD). However, using a 70-ms cutoff criterion produced better specificity (i.e., 75% of MTD cases were correctly classified). The tradeoff between sensitivity and specificity

Table 1. Mean number of breaks of ADSD versus MTD (including men and women and adjusted for gender). P *values reflect the result of independent samples* t *tests comparing group means.*

Duration of Breaks	Mean Number of Phonatory Breaks	Standard Error	*t* Statistic	*p* value
20 ms				
MTD	0.82	0.38	-2.67	0.0088
ADSD	2.26	0.39		
30 ms				
MTD	0.70	0.37	-2.59	0.0109
ADSD	2.07	0.38		
40 ms				
MTD	0.50	0.30	-2.93	0.0042
ADSD	1.73	0.31		
50 ms				
MTD	0.33	0.27	-2.93	0.0042
ADSD	1.46	0.28		
60 ms				
MTD	0.25	0.26	-2.68	0.0086
ADSD	1.25	0.27		
70 ms				
MTD	0.27	0.23	-1.92	0.0584
ADSD	0.90	0.24		
80 ms				
MTD	0.25	0.20	-1.58	0.1173
ADSD	0.69	0.20		
90 ms				
MTD	0.22	0.17	-1.70	0.0923
ADSD	0.62	0.17		
100 ms				
MTD	0.20	0.16	-1.53	0.1285
ADSD	0.55	0.17		

Table 2. Percentage of male participants in ADSD versus MTD according to phonatory breaks at durations 40-70 ms.

	≥40 ms	≥50 ms	≥60 ms	≥70 ms
ADSD	63%	42%	32%	32%
MTD	20%	10%	10%	0%
X^2	4.89	3.15	1.67	3.98
p value	0.0271	0.0757	0.1968	0.0460

Table 3. Percentage of female participants in ADSD versus MTD according to phonatory breaks at durations 40-70 ms.

	≥40 ms	≥50 ms	≥60 ms	≥70 ms
ADSD	68%	64%	64%	50%
MTD	47%	45%	39%	31%
X^2	2.75	2.13	3.77	2.46
p value	0.0970	0.1442	0.0521	0.1169

Table 4. Estimates of sensitivity, specificity, PV+, PV-, LR+, & LR- for ADSD according to duration (including men and women).

Duration	Sensitivity	Specificity	PV+	PV-	LR+ (95% CI)	LR- (95% CI)
20 ms	0.71	0.41	0.45	0.67	0.20 (0.89, 1.59)	0.71 (0.40, 1.27)
30 ms	0.68	0.47	0.47	0.68	0.28 (0.94, 1.79)	0.68 (0.40, 1.13)
40 ms	0.66	0.58	0.52	0.71	0.57 (1.07, 2.25)	0.59 (0.37, 0.96)
50 ms	0.54	0.61	0.49	0.65	0.38 (0.90, 2.11)	0.75 (0.52, 1.12)
60 ms	0.49	0.66	0.50	0.65	0.44 (0.90, 2.31)	0.77 (0.55, 1.10)
70 ms	0.41	0.75	0.53	0.65	0.64 (0.92, 2.88)	0.79 (0.58, 1.06)
80 ms	0.34	0.78	0.52	0.63	0.55 (0.82, 2.94)	0.85 (0.65, 1.09)
90 ms	0.32	0.78	0.50	0.62	0.45 (0.75, 2.78)	0.87 (0.68, 1.12)
100 ms	0.29	0.80	0.50	0.62	0.45 (0.72, 2.88)	0.89 (0.70, 1.12)

will be discussed in more detail later. Based upon inspection of the PV+ and PV- only, it is clear that duration has only a modest influence on these values which ranged from 0.45 to 0.52 for PV+, and 0.62 to 0.71 for PV-. However, the 40-ms cutoff seems to optimize both PV+ and PV-, providing the most precise diagnostic information on the basis of this analysis. For instance, at a 40-ms cutoff, a negative test (i.e., no breaks of this duration) produced a probability of 0.71 that the patient did not have ADSD, whereas a positive test resulted in a probability of 0.52 that the patient did have ADSD. However, the LR+ and LR- results for both men and women failed to produce conclusive results.

Tables 5 and 6 provide the same diagnostic indices as noted above, but data reported separately for males and females. This separation provides improved resolution, and reveals remarkable differences in specificity, PV+, PV-, LR+, and LR-, particularly

for males. For males, at greater or equal to 70-ms, estimates of specificity become 1.00, indicating that no male subjects with MTD had phonatory breaks in excess of 70-ms. Similarly for males, PV+ also reach 1.00 indicating perfect diagnostic precision such that a positive test result (i.e., the finding of a diagnostic break greater than 70-ms) confirms that the patient has ADSD. Furthermore, LR+ increased to 6.50, indicating moderate to high odds that an individual will have ADSD when presenting with phonatory breaks at this duration. In contrast, inspection of the results obtained from the female analysis does not show comparable separation of the groups leading to reliable differentiation of ADSD or MTD (Table 6). Thus, separating males and females shows an interesting phenomenon in the ability of the diagnostic test (i.e., phonatory break analysis) to differentiate between ADSD and MTD. Males are more easily distinguished as having ADSD or MTD than females.

Although phonatory breaks are more characteristic of ADSD, they can still occur in patients with MTD (especially females). Figures 3 and 4 provide examples of phonatory breaks on the word *avenue* in a male participant with ADSD and a female participant with MTD. The duration of the phonatory breaks in the two figures are both larger than the 40-ms cutoff identified but the number of phonatory breaks within the sample should also be considered when using phonatory break analysis as a diagnostic tool.

The diagnostic performance of combining information regarding both duration and number of phonatory breaks is explored in Table 7. Inspection of sensitivity, specificity, PV+, and PV- at 40, 50, and 60-ms intervals demonstrates greater specificity if there is more than one phonatory break (see Table 7). Thus, the addition of

Table 5. Estimates of sensitivity, specificity, PV+, PV-, LR+, & LR- according to duration of phonatory breaks for males only.

Duration	Sensitivity	Specificity	PV+	PV-	LR+ (95% CI)	LR- (95% CI)
20 ms	0.68	0.70	0.81	0.54	2.28 (0.84, 6.17)	0.45 (0.21, 0.98)
30 ms	0.68	0.70	0.81	0.54	2.28 (0.84, 6.18)	0.45 (0.21, 0.99)
40 ms	0.63	0.80	0.86	0.53	3.16 (0.87, 11.43)	0.46 (0.24, 0.90)
50 ms	0.42	0.90	0.89	0.45	4.21 (0.61, 29.09)	0.64 (0.42, 0.99)
60 ms	0.42	0.90	0.86	0.41	3.16 (0.44, 22.73)	0.76 (0.53, 1.10)
70 ms	0.32	1.00	1.00	0.43	6.50 (0.41, 104.20)	0.68 (0.50, 0.93)
80 ms	0.26	1.00	1.00	0.42	6.05 (0.37, 99.49)	0.74 (0.56, 0.96)
90 ms	0.26	1.00	1.00	0.42	7.05 (0.37, 99.50)	0.74 (0.56, 0.97)
100 ms	0.21	1.00	1.00	0.40	4.95 (0.29, 83.68)	0.79 (0.63, 0.99)

Table 6. Estimates of sensitivity, specificity, PV+, PV-, LR+, & LR- according to duration of phonatory breaks for females only.

Duration	Sensitivity	Specificity	PV+	PV-	LR+ (95% CI)	LR- (95% CI)
20 ms	0.73	0.35	0.33	0.74	1.11 (0.80, 1.54)	0.79 (0.36, 1.72)
30 ms	0.68	0.43	0.35	0.75	1.19 (0.82, 1.74)	0.74 (0.37, 1.48)
40 ms	0.68	0.53	0.39	0.79	1.45 (0.96, 2.19)	0.60 (0.31, 1.17)
50 ms	0.64	0.55	0.39	0.77	1.42 (0.91, 2.21)	0.66 (0.36, 1.21)
60 ms	0.64	0.61	0.42	0.79	1.64 (1.02, 2.63)	0.59 (0.32, 1.08)
70 ms	0.50	0.69	0.42	0.76	1.63 (0.90, 2.96)	0.72 (0.46, 1.14)
80 ms	0.41	0.73	0.41	0.73	1.54 (0.78, 3.06)	0.80 (0.55, 1.18)
90 ms	0.36	0.73	0.38	0.72	1.37 (0.67, 2.82)	0.87 (0.61, 1.24)
100 ms	0.36	0.76	0.40	0.73	1.48 (0.71, 3.11)	0.84 (0.59, 1.20)

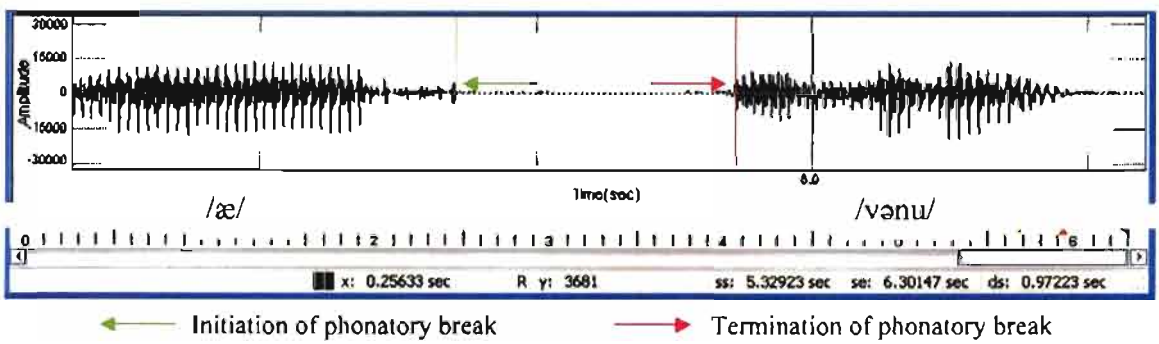


Figure 3. Phonatory break from ADSD patient on the word avenue *measuring 256.33 ms*

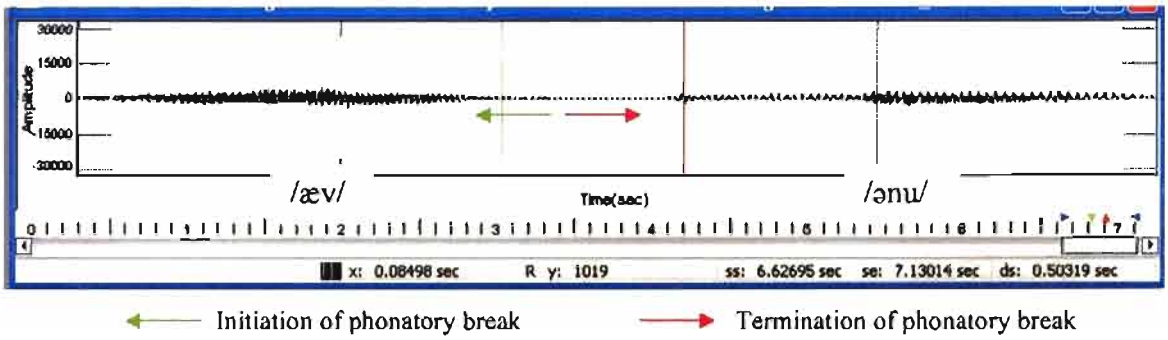


Figure 4. Phonatory break from MTD patient on the word avenue *measuring 84.98-ms*

information regarding the frequency of phonatory breaks appears to result in improved diagnostic precision over duration information alone. As the number of phonatory breaks increases, there is a significant return in specificity estimates at every duration such that specificity estimates exceed 80% if three or more breaks occur, indicating that patients with MTD rarely exhibit more than two or three breaks, regardless of the duration of these breaks. Furthermore, this relationship is highlighted in the 4+ break category which produces the highest PV+, PV-, LR+, and LR-, suggesting that not only does the duration of phonatory break influence diagnostic precision, but the number of breaks enhances the performance of phonatory break analysis in differential diagnosis.

In Tables 8 and 9 data for males and females are presented separately specifically for phonatory breaks durations between 40-60 ms. These data separated by gender illustrate the improved precision of the diagnostic test. In this case, adding information regarding the frequency of phonatory breaks observed increases the likelihood of correct classification/diagnosis. This is particularly evident in Table 9 when four or more breaks are present (regardless of the duration) which improves all indices of diagnostic precision for the female group.

Table 7. Sensitivity, specificity, PV+, PV-, LR+, & LR- for ADSD according to selected phonatory break points and comparisons of zero versus X number of breaks (including men and women).

Duration	Breaks	Sens.*	Spec.†	PV+	PV-	LR+ (95% CI)	LR- (95% CI)
40 ms							
	0 vs. 1+	0.61	0.60	0.49	0.71	1.53 (1.01, 2.28)	0.65 (0.41, 1.31)
	0 vs. 2+	0.46	0.74	0.50	0.71	1.77 (0.93, 3.35)	0.73 (0.49, 1.08)
	0 vs. 3+	0.36	0.83	0.53	0.71	2.12 (0.89, 5.09)	0.77 (0.54, 1.08)
	0 vs. 4+	0.33	0.94	0.78	0.71	5.50 (1.37, 26.26)	0.71 (0.52, 0.97)
50 ms							
	0 vs. 1+	0.54	0.61	0.49	0.65	1.38 (0.90, 2.11)	0.75 (0.52, 1.12)
	0 vs. 2+	0.39	0.75	0.50	0.65	1.56 (0.80, 3.00)	0.81 (0.59, 1.13)
	0 vs. 3+	0.30	0.84	0.53	0.65	1.88 (0.75, 4.45)	0.83 (0.64, 1.11)
	0 vs. 4+	0.27	0.95	0.78	0.65	5.40 (1.15, 22.70)	0.77 (0.60, 0.99)
60 ms							
	0 vs. 1+	0.49	0.66	0.50	0.65	1.44 (0.90, 2.31)	0.77 (0.55, 1.10)
	0 vs. 2+	0.36	0.76	0.50	0.65	1.50 (0.79, 3.02)	0.84 (0.62, 1.12)
	0 vs. 3+	0.28	0.85	0.53	0.65	1.87 (0.74, 4.47)	0.85 (0.66, 1.10)
	0 vs. 4+	0.25	0.95	0.78	0.65	5.00 (1.15, 22.88)	0.79 (0.62, 0.99)

* = sensitivity † = specificity

ROC Analysis

Using sensitivity and 1-specificity, values were used to create an ROC curve. Inspection of the ROC curve combined for males and females confirms that 40-ms is the ideal duration for determining diagnostic precision. The 40-ms duration offers 66% sensitivity and 58% specificity in identifying ADSD from MTD. As stated before, this duration also shows respectable PV- indicating that those who do not have phonatory breaks will be correctly identified 71% of the time as not having ADSD. A LR+ of 0.57 and LR- of 0.59 with males and females combined indicate less impressive values for separating the two disorders. Figure 5 is the graphical representation of the ROC curve. Cutoff values used were the intervals of the duration of the phonatory breaks for ADSD and MTD.

Table 8. Sensitivity, specificity, PV+, PV-, LR+, & LR- according to selected phonatory break durations and comparisons of zero versus X number of breaks for males.

Duration	Breaks	Sens.*	Spec.†	PV+	PV-	LR+ (95% CI)	LR- (95% CI)
40 ms							
	0 vs. 1+	0.53	0.89	0.89	0.53	4.80 (0.71, 32.34)	0.53 (0.29, 0.95)
	0 vs. 2+	0.22	1.00	1.00	0.53	4.50 (0.25, 81.76)	0.78 (0.55, 1.10)
	0 vs. 3+	0.13	1.00	1.00	0.53	3.00 (0.14, 64.26)	0.88 (0.67, 1.14)
	0 vs. 4+	0.00	1.00	1.00	0.53	-	-
50 ms							
	0 vs. 1+	0.42	0.90	0.89	0.45	4.21 (0.61, 29.09)	0.64 (0.42, 0.99)
	0 vs. 2+	0.15	1.00	1.00	0.45	3.57 (0.19, 66.61)	0.85 (0.67, 1.07)
	0 vs. 3+	0.08	1.00	1.00	0.45	2.31 (0.10, 50, 85)	0.92 (0.77, 1.09)
	0 vs. 4+	0.00	1.00	-	0.45	-	-
60 ms							
	0 vs. 1+	0.32	0.90	0.86	0.41	3.16 (0.44, 22.73)	0.76 (0.53, 1.10)
	0 vs. 2+	0.13	1.00	1.00	0.41	3.13 (0.17, 58.63)	0.87 (0.71, 1.06)
	0 vs. 3+	0.07	1.00	1.00	0.41	2.00 (0.09, 44.35)	0.93 (0.80, 1.07)
	0 vs. 4+	0.00	1.00	-	0.41	-	-

* = sensitivity † = specificity *Not estimable because of small numbers.

Table 9. Sensitivity, specificity, PV+, PV-, LR+, & LR- according to selected phonatory break durations and comparisons of zero versus X number of breaks for females.

Duration	Breaks	Sens.*	Spec.†	PV+	PV-	LR+ (95% CI)	LR- (95% CI)
40 ms							
	0 vs. 1+	0.67	0.54	0.39	0.79	1.45 (0.94, 2.24)	0.62 (0.32, 1.19)
	0 vs. 2+	0.59	0.68	0.45	0.79	0.86 (1.01, 3.44)	0.60 (0.33, 1.11)
	0 vs. 3+	0.50	0.79	0.50	0.79	2.36 (1.02, 5.46)	0.63 (0.37,1.10)
	0 vs. 4+	0.50	0.93	0.78	0.79	7.00 (1.67, 29.38)	0.54 (0.32, 0.92)
50 ms							
	0 vs. 1+	0.64	0.55	0.39	0.77	1.42 (0.91, 2.21)	0.66 (0.36, 1.21)
	0 vs. 2+	0.56	0.69	0.45	0.77	1.81 (0.97, 3.38)	0.64 (0.37, 1.12)
	0 vs. 3+	0.47	0.79	0.50	0.77	2.27 (0.97, 5.32)	0.67 (0.41, 1.11)
	0 vs. 4+	0.47	0.93	0.78	0.77	6.77 (1.60, 28.63)	0.57 (0.35, 0.93)
60 ms							
	0 vs. 1+	0.64	0.61	0.42	0.79	1.64 (1.02, 2.63)	0.59 (0.33, 1.08)
	0 vs. 2+	0.56	0.71	0.45	0.79	1.94 (1.03, 3.66)	0.62 (0.36, 1.08)
	0 vs. 3+	0.47	0.81	0.50	0.79	2.47 (1.05, 5.82)	0.66 (0.40, 1.08)
	0 vs. 4+	0.47	0.94	0.78	0.79	7.47 (1.76, 31.73)	0.57 (0.35, 0.92)

* = sensitivity † = specificity *Not estimable because of small numbers.

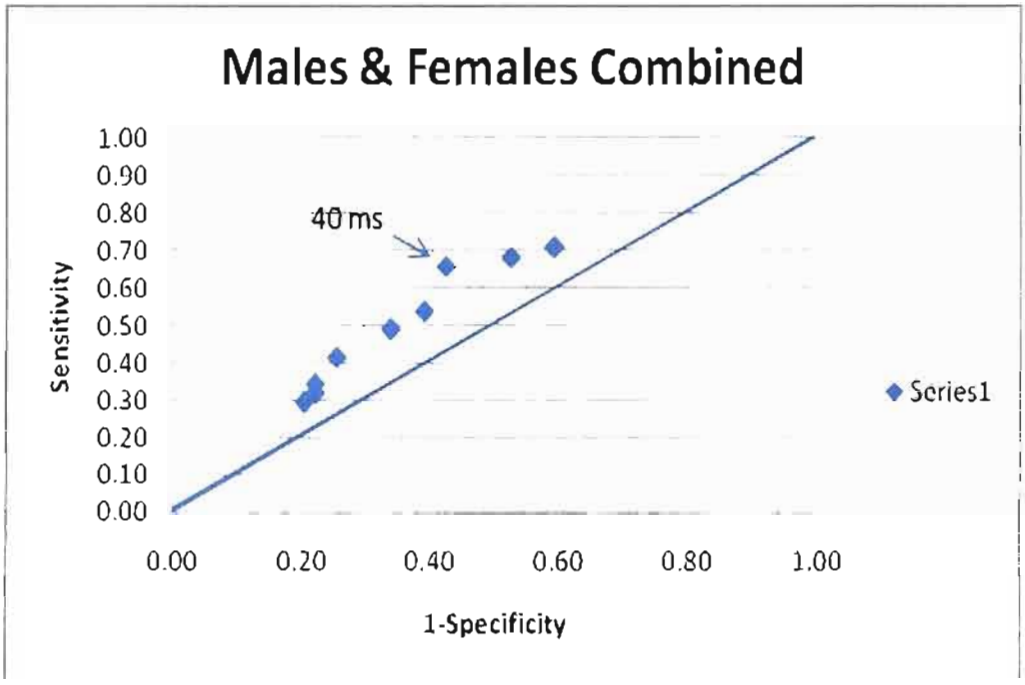


Figure 5. ROC curve of ADSD and MTD participants (including men and women)

DISCUSSION

The purpose of this study was to examine the diagnostic precision and utility of phonatory break analysis in the differential diagnosis of ADSD and MTD. It has been asserted that phonatory breaks are sine qua non of ADSD, but little objective acoustic evidence exists to substantiate this claim. Furthermore, it is unknown whether individuals with MTD also demonstrate phonatory breaks. In this study, the total number and duration of phonatory breaks were acoustically determined for subjects with confirmed ADSD and MTD during production of an all-voiced sentence. The results revealed that (a) individuals with ADSD showed a higher mean number of phonatory breaks as compared to individuals with MTD; (b) estimates of sensitivity and specificity, along with other measures of diagnostic precision, varied according to both duration and frequency of phonatory breaks; (c) males and females with MTD differed with respect to the duration and frequency of phonatory breaks, leading to very different test performance results; and (d) combining information regarding duration and frequency, along with knowledge of gender of the patient improved diagnostic test performance. In the following paragraphs these conclusions are discussed more fully, and interpreted within the context of previous research in this area.

When male and female data were separated—according to duration—notable changes occurred in estimates of sensitivity, specificity, PV+, PV-, LR+, and LR-.

Interestingly, the results indicated that diagnostic precision for male subjects as compared to females improved substantially, as it is apparent that male subjects with MTD did not present with phonatory breaks at the same frequency as their female counterparts. Inspection of the results confirmed that males with MTD rarely evidenced phonatory breaks, and no subject with MTD had a break greater than 70-ms. Therefore, all measures of diagnostic precision were markedly better in males as compared to females, suggesting that phonatory break analysis is a useful diagnostic test especially in males. This separation is particularly relevant considering that in the Sapienza et al. (2000) study which compared subjects with MTD and ADSD, no distinction was made between males and females.

In conjunction with the duration of phonatory breaks, knowledge regarding number of phonatory breaks appeared to provide additional valuable discriminatory information regarding characteristics of ADSD and MTD, especially for females. As the number of phonatory breaks increases, diagnostic precision also increases. For example, when a patient has more than four phonatory breaks, it can be quite confidently concluded that the patient has ADSD as evidenced by the highest specificity, PV+, PV-, LR+, and LR- values being associated with greater number of phonatory breaks. According to these diagnostic indices, patients with MTD (regardless of gender) will almost never have more than four phonatory breaks in the single all-voiced sentence. Therefore, combining information regarding both the duration and the frequency of phonatory breaks affords greater diagnostic precision in discriminating ADSD and MTD. Combining all results as discussed above provides three factors that contribute to the

highest diagnostic precision for ADSD and MTD: (a) duration of phonatory breaks; (b) number of phonatory breaks; and (c) gender.

These results differ in substantial ways with the results reported by Sapienza et al. (2000) and Rees et al. (2007). Both of these studies reportedly failed to find phonatory breaks in any of their participants with MTD. In contrast, results of the present study confirmed that the group with ADSD showed significantly more phonatory breaks as compared to the MTD group, but phonatory breaks were not exclusively the domain of ADSD, as almost 50% of the female participants with MTD also showed evidence of phonatory breaks. This difference likely reflects differences in exclusion criteria, especially with Sapienza et al., who excluded anyone with MTD who showed perceptual evidence of phonatory breaks. Second, although Sapienza et al. used a 50-ms break duration as the criterion for counting phonatory breaks, this study explored all phonatory breaks greater than 20-ms. When looking at the number of phonatory breaks at various durations in ADSD and MTD overall, it is clear that individuals with ADSD demonstrate more frequent phonatory breaks as compared to individuals with MTD at durations less than 50-ms. This was particularly evident with breaks which ranged from 20-60 ms. Therefore, it can be concluded that patients with ADSD (combined and adjusted for males and females) have more phonatory breaks than patients with MTD, particularly at durations less than 60-ms. The ideal duration of phonatory break measurement is 40-ms as evidenced by estimates of sensitivity, specificity, PV+, PV-, LR+, and LR-.

Finally, it is clear that different patterns exist for men and women with MTD, which were not addressed by Sapienza et al. (2000) or Rees et al. (2007). Men with MTD rarely exhibited phonatory breaks, and as such the diagnostic precision of phonatory

break analysis was much improved when applied to male participants only. In light of this difference, it can be concluded that the presence of phonatory breaks in males sufficiently distinguishes ADSD from MTD. However, phonatory break analysis as a diagnostic test is less precise when applied to women.

One impediment to interpreting previous studies comparing the value of phonatory break analysis in distinguishing ADSD and MTD is the lack of estimates of sensitivity and specificity. This study included both of these measures, along with other measures of diagnostic precision to increase clinical utility. From these results, it appears for both genders combined, the optimal break duration proved to be 40-ms as evidenced by 66% sensitivity, 52% PV+, and 71% PV-. Therefore, duration plays an influential role in determining diagnosis of ADSD and MTD and should be more closely examined when performing diagnostic tests.

Clinical Utility

Based upon the diagnostic indices reported here, phonatory break analysis provides some promise as a means to distinguish ADSD from MTD, especially in males. One limitation of phonatory break analysis is the time-consuming nature of manually identifying and measuring phonatory breaks. Given the encouraging results reported here, an automated process should be explored to permit faster acoustic analysis of phonatory breaks of speech samples collected from patients. Furthermore, additional research is required to understand other acoustic features of ADSD and MTD to better differentiate the two disorders.

Certainly, combining the three factors described above including duration of phonatory breaks, frequency of phonatory breaks, and gender produces improved yet

imperfect, diagnostic resolution. Even combining these factors, there is still some likelihood of misdiagnosis due to imperfect knowledge about the disorders and their individual characteristics. Therefore, it is still advisable to rely on a composite from the results of other diagnostic tasks revealing task-specific performance in ADSD (Roy et al., 2005, 2007). These include differences in sustained vowel production versus connected speech, or voiced versus voiceless-laden sentences, which maximally provoking sign expression in ADSD.

CONCLUSION

In addition to differentiating between ADSD and MTD perceptually, auditory voice evaluation can assist in helping increase diagnostic accuracy. This investigation produced results considering the number and duration of phonatory breaks in the two disorders, gender of the speaker, and how phonatory breaks correlate with perceptual severity ratings. It can be concluded that individuals with ADSD produce more phonatory breaks than individuals with MTD. Furthermore data indicate that phonatory break analysis is more sensitive in differentiating between ADSD and MTD for males. Acoustic phonatory break analysis can be combined with perceptual severity ratings to provide even stronger evidence to support the distinction between ADSD and MTD.

REFERENCES

Aronson, A. E. (1990). *Clinical voice disorders: An interdisciplinary approach* (3rd ed.). New York: Thieme.

Cannito, M. P., & Kondraske, G. V. (1990). Rapid manual abilities in spasmodic dysphonia and normal female subjects. *Journal of Speech and Hearing Research, 33,* 123–133.

Cannito, M. P., & Woodson, G. (2000). The spasmodic dysphonias. In R. Kent, & M. Ball (Eds.), *Voice quality measurement.* San Diego: Singular Thomson Learning.

Cimino-Knight, A. M., & Sapienza, C. M. (2001). Consistency of voice produced by patients with adductor spasmodic dysphonia: A preliminary investigation. *Journal of Speech, Language, and Hearing Research, 44,* 793–802.

Dedo, H. H., & Shipp, T. (1980). *Spastic dysphonia: A surgical and voice therapy treatment program.* Houston: College Hill Press.

Dollaghan, C. A. (2007). *The handbook for evidence-based practice in communication disorders.* Baltimore: Brookes.

Fairbanks, G. (1960). *Voice and articulation drillbook* (2nd ed). New York: Harper & Row, 124-139.

Gouse, M. L. (2004). *Muscle tension dysphonia versus adductor spasmodic dysphonia: Exploring the influence of voice context on the severity of voice symptoms.* Unpublished master's thesis, University of Utah, Salt Lake City.

Higgins, M. B., Chait, D. H., & Schulte, L. (1999). Phonatory air flow characteristics of adductor spasmodic dysphonia and muscle tension dysphonia. *Journal of Speech, Language, and Hearing Research, 42,* 101–111.

Hillenbrand, J. M., & Gayvert, R. T. (2005). Open source software for experiment design and control. *Journal of Speech, Language, and Hearing Research, 48,* 45-55.

Houtz, D. (2006). *Muscle tension dysphonia versus adductor spasmodic dysphonia: Exploring the influence of voice context on the severity of voice symptoms.* Unpublished master's thesis, University of Utah, Salt Lake City.

Langeveld, T. P. M., Drost, H. A., Frijns, J. H. M., Zwinderman, A. H., & Baatenburg de Jong, R. J. (2000). Perceptual characteristics of adductor spasmodic dysphonia. *The Annals of Otology, Rhinology, and Laryngology, 11,* 214–222.

Leonard, R., & Kendall, K. (1999). Differentiation of spasmodic and psychogenic dysphonias with phonoscopic evaluation. *Laryngoscope, 109*(2, part 1), 295–300.

Ludlow, C. L. (1995). *Management of the spasmodic dysphonias.* In J. S. Reubin, R. T. Sataloff, G. Korovin, W. J. Gould, (Eds.), *Diagnosis and treatment of voice disorders.* New York: Igaku-Shoin, 436-454.

Ludlow, C. L., & Mann, E. A. (2003). From the 25th annual Paul Streit Memorial Seminar. *Laryngology Symposium-Evaluation and Management of Voice Disorders.* Walter Reed Army Medical Center Otolaryngology-Head and Neck Surgery Service, 134-217.

Rees, C. J., Blalock, P. D., Kemp, S. E., Halum, S. L., & Koufman, J. A. (2007). Differentiation of adductor-type spasmodic dysphonia from muscle tension dysphonia by spectral analysis. *Otolaryngology—Head and Neck Surgery, 137,* 576-581.

Roy, N., Bless, D. M., Heisey, D., & Ford, C. N. (1997). Manual circumlaryngeal therapy for functional dysphonia: An evaluation of short- and long-term treatment outcomes. *Journal of Voice, 11,* 321–331.

Roy, N., Gouse, M., Mauszycki, S. C., Merrill, R. M., & Smith, M. E. (2005). Task specificity in adductor spasmodic dysphonia versus muscle tension dysphonia. *The Laryngoscope, 115,* 311–316.

Roy, N., & Leeper, H. A. (1993). Effects of the manual laryngeal musculoskeletal tension reduction technique as a treatment for functional voice disorders: Perceptual and acoustic measures. *Journal of Voice, 7,* 242–249.

Roy, N., Mauszycki, S. C., Merrill, R. M., Gouse, M., & Smith, M. E. (2007). Toward improved differential diagnosis of adductor spasmodic dysphonia and muscle tension dysphonia. *Folia Phoniatrica & Logopedica, 59,* 83–90.

Roy, N., McGrory, J. J., Tasko, S. M., Bless, D. M., Heisey, D., & Ford, C. D. (1997). Psychological correlates of functional dysphonia: An investigation using the Minnesota Multiphasic Personality Inventory. *Journal of Voice, 11,* 443–451.

Roy, N., Smith, M. E., Allen, B., & Merrill, R. M. (2007). Adductor spasmodic dysphonia versus muscle tension dysphonia: Examining the diagnostic value of recurrent laryngeal nerve lidocaine block. *Annals of Otology, Rhinology, and Laryngology, 116*(3), 161–168.

Sapienza, C. M., Murry, T., & Brown, W. S. Jr. (1998). Variations in adductor spasmodic dysphonia: Acoustic evidence. *Journal of Voice, 12,* 214–222.

Sapienza, C. M., Walton, S., & Murry, T. (1999). Acoustic variations in adductor spasmodic dysphonia as a function of speech task. *Journal of Speech, Language, and Hearing Research, 42,* 127–140.

Sapienza, C. M., Walton, S., & Murry, T. (2000). Adductor spasmodic dysphonia and muscular tension dysphonia: Acoustic analysis of sustained phonation and reading. *Journal of Voice, 14,* 502–520.